THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF COMPUTING

Volume 61 November 1982 Number 9, Part 2

Database Systems:

Database Work at Bell Laboratories

By A. V. AHO

(Manuscript received February 23, 1982)

This computing issue of The Bell System Technical Journal is devoted to papers discussing some of the many database applications at Bell Laboratories. This paper introduces the themes described more fully in the other papers in the issue.

I. INTRODUCTION

There is no easy way to characterize the many different kinds of data that are used to run the Bell System. The volume of data involved is itself staggering. At Bell Laboratories alone, more than two-trillion bytes are stored on various kinds of on-line devices in the computer centers. That's equivalent to more than one-million good-sized textbooks. Clearly, computerized systems are necessary to administer and keep track of such large quantities of data. But because the uses of data are so varied, no one data management system has proven suitable for all applications.

The papers in this issue are a sampling of the broad range of database applications at Bell Laboratories. The papers were presented at a database symposium held at Bell Laboratories in April 1981. Although great diversity in application areas is evident, many of the issues discussed in these papers are common to all database systems-development efforts. It is hoped that by exposing these issues here,

progress towards a more unified approach for database systems design will be hastened.

II. DATABASE SYSTEMS WITH THE *UNIX** OPERATING SYSTEM

The design of database management systems as part of the *UNIX* operating system, and the evaluation of their performance, is still a lively research activity. The first paper by M. J. Rochkind describes a flexible database file system constructed on top of the *UNIX* operating system. This file system provides primitive functions out of which database management systems can be built.

Although more than a dozen commercial database management systems have been built around the *UNIX* operating system, there is still some question on how to achieve maximum performance from such systems. In the second paper, P. J. Weinberger discusses some of the relevant underlying parameters and suggests simple but effective methods by which the *UNIX* operating system can be modified to suit database applications better.

III. DBMS FOR BELL SYSTEM APPLICATIONS

The next four papers describe database management systems that have been built for specific Bell System applications. The paper by Barclay, Byrne, and Ng describes a database manager that provides the interface between the application programs and the data in the No. 5 Electronic Switching System, the newest member of the ESS family. Good real-time response was an important design goal for this database manager.

Then C. C. Wang and C. P. Huang describe DBAS, a relational database management system written in the C programming language to run under the *UNIX* operating system. This system is used to provide administrative and service-order support for operating telephone companies. High volume performance is critical: the database system needs to handle databases of millions of records and transaction rates approaching 100,000 updates per day. The paper by T. C. Chiang and G. R. Rose describes DMB-2, a database management system that supports the entity-relationship data model. This system is used to track repair operations in a telephone company. The design goals included data independence and ease of change without unduly sacrificing performance.

Modeling real-world data in a way that can be effectively used in a database environment is a central problem of database design. Providing a working data model for local telephone loop plant (this consists

* Trademark of Bell Laboratories.

of the wires and other equipment that connect a central office to a telephone customer's premises) was long considered a Gordian knot. A. J. Goldstein in his paper describes an elegant and practical solution to this problem using a directed hypergraph data model.

IV. DISTRIBUTED SYSTEMS

Distributed database problems arise naturally in a number of Bell System applications. David Cohen describes the implementation of an effective multiple-copy update algorithm that was used in a distributed data management system to support logical networks.

The paper by J. P. Linderman describes the design of a distributed record management system in which the emphasis is easy interchangeability of components. The following paper by W. D. Roome describes a content-addressable intelligent store that services the record manager in this distributed record management system. The intelligent store uses a parallel search disk to provide real-time read-write access to pages of information.

V. INFORMATION RETRIEVAL AND HUMAN FACTORS

The paper by Nancy Goguen describes a new information-services system that combines several distinct databases to provide a broad range of information retrieval services. Indexed retrieval, browsing, and customization of information are among the services offered by the system.

The last paper, by T. K. Landauer and his colleagues, addresses an issue that is becoming even more important in database systems design. This paper discusses the implications of human factors in data access and in the design of information retrieval systems.

As we have mentioned, these papers represent only a sampling of the broad spectrum of database work at Bell Laboratories. However, in spite of great diversity, certain unifying themes are starting to emerge. The importance of data models, modular architecture, performance, and manageable software components is already evident. And perhaps, most significantly, future database systems will continue the evolution toward friendlier user interfaces.